

Achievable speedup in CUDA with hand-tuned GPU data transfer and execution overlap

